DOCUMENT RESUME

ED 130 611 IR 004 139

AUTHOR	Swets, John A.; And Others
TITLE	Learning to Identify Nonverbal Sounds: An Application of a Computer as a Teaching Machine. Technical Report.
INSTITUTION	Bolt, Beranek and Newman, Inc., Cambridge, Mass.
SPONS AGENCY	Naval Training Device Center, Orlando, Fla.
REPORT NO	NAVTRADEVCEN-789-1
PUB DATE	23 Apr 62
CONTRACT	N-61339-789
NOTE	49p.; Archival document
EDRS PRICE	MF-$0.83 HC-$2.06 Plus Postage.
DESCRIPTORS	*Aural Learning; College Students; *Computer Assisted Instruction; Educational Research; Feedback; Overt Response
IDENTIFIERS	Learner Controlled Instruction

ABSTRACT

To test subject's capacity for auditory learning, four experiments were conducted in which college students were given the task of identifying specific multidimensional sounds after a period of practice. A computer-based teaching machine was used which incorporated the following features: (1) continual interrogation and overt response; (2) immediate feedback or knowledge of results; (3) learner-controlled pacing; and (4) presentation of successive items conditioned on previous performance. Experiments attempted to maximize performance by manipulating learning variables, and results proved to be similar to those obtained previously using conventional teaching methods. Certain of the central features of automated instruction were found to hinder learning in the tasks studied. This report briefly describes procedures for the experiments, and results are summarized. (EMH)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED130611

Technical Report: NAVTRADEVCEN 789-1

LEARNING TO IDENTIFY NONVERBAL SOUNDS : AN APPLICATION OF A COMPUTER AS A TEACHING MACHINE

John A. Swets

with the collaboration of

Susan H. Millman
William E. Fletcher
David M. Green
J. C. R. Licklider
Edward Fredkin
Earl E. Winter

U S DEPARTMENT OF HEALTH
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

23 April 1962

Bolt, Beranek and Newman, Inc.
Cambridge, Massachusetts

U. S. NAVAL TRAINING DEVICE CENTER
Port Washington, New York

Contract No. N61339-789

IR 000 139

LEARNING TO IDENTIFY NONVERBAL SOUNDS: AN APPLICATION OF A COMPUTER AS A TEACHING MACHINE

ABSTRACT

The procedures of automated instruction -- continual interrogation and overt response, immediate knowledge of results, presentation of successive items conditional upon previous performance, learner-controlled pacing of the lesson, and so forth -- were applied to the task of learning to identify multidimensional, nonverbal sounds. These procedures produced results that are comparable to those obtained previously with conventional training methods. Certain of the central features of automated instruction were found to hinder learning in the task studied.

FOREWORD

Purpose

Teaching machine or programed instruction techniques have engendered widespread interest in educational circles. These techniques have potentially important implications for military training situations wherein the time available for training is short and the student input relatively heterogeneous. The U. S. Naval Training Device Center is supporting a research program directed toward examining teaching machine variables and their military application. The research described in this report is part of that program.

The purpose of the study reported here is to explore the utility of generally accepted principles of programed instruction in developing non-verbal skills -- specifically sound identification. This area of study is particularly relevant to military training because a substantial number of the skill complexes required in military operations are essentially non-verbal.

Results

Four experiments were conducted in the area of auditory learning. A series of multidimensional sounds was presented to college students acting as subjects. Their task was to identify certain of these sounds after a period of practice. Various techniques of programed instruction were used during the practice periods and compared with conventional training methods.

The results of these experiments are described in the Brief of the Study (pg. 2). In general, those groups trained by means of teaching machine principles were not superior to those trained in conventional ways, and in certain instances were actually inferior, in identifying non-verbal sounds.

Implications

Use should be made of automated training procedures in developing auditory skills only after a careful investigation of their efficiency. In fact, at this stage of development,

great care should be exercised in the application of these principles to any non-verbal training situation. Continuing research is directed toward allowing the subject greater freedom in manipulating the training situation and in examining other kinds of non-verbal behavior such as visual perception.

James J. Regan, Ph.D.
Head, Systems Psychology Division

K. F. Thomson, Ph.D.
Chief Psychologist

TABLE OF CONTENTS

	Page
RECOMMENDATIONS	1
BRIEF OF STUDY	2
REFERENCES	25
TECHNICAL APPENDICES	
Appendix A - Description of Computer	26
Appendix B - Programming and Lesson Procedures	31
Appendix C - Sound Parameters	39
Appendix D - Research Recommendations	40

LIST OF TABLES

Table		Page
1	The Results of Experiment III	17

LIST OF ILLUSTRATIONS

Figure		
1	The Results of Experiment I.....	9
2	The Results of Experiment II	15
3	The Results of Experiment IV	20
A-1	A Photograph of the PDP-1B Computer	30
B-1	A Flow Chart of an Illustrative Computer Program	32

RECOMMENDATIONS

This is a report of a laboratory study, and the majority of the recommendations are made with respect to further research. These are presented in Appendix D, p. 40. As concerns practice, it can be strongly recommended, on the basis of this research, that the institution of automated training procedures for teaching identification of auditory displays follow only upon a very careful investigation of the efficiency of these procedures. An unconsidered application of these procedures, which is certainly natural enough in the light of the enthusiasm and hopes that they have engendered, may actually be detrimental in the case of auditory learning.

BRIEF OF STUDY

Introduction

When we attempt to identify stimuli that differ in only one respect, e.g., tones that differ in frequency, we can identify them correctly if there are no more than six or seven alternatives. If there are more alternatives, we make mistakes -- enough mistakes so that, however many stimulus alternatives are presented, our performance is equivalent to errorless identification of six or seven stimuli. Our one-dimensional judgments, in other words, are capable of transmitting approximately $\log_2 6$, or 2.6 bits of information. In making this observation, George Miller summarized the results of several studies which included judgments of auditory, visual, gustatory, and tactile stimuli.[1] The figure of 2.6 bits is the mean result; the range is from 1.6 bits, or three alternatives, to 3.9 bits, or 15 alternatives.

Miller also reviewed the results of attempts to identify stimuli that vary along more than one dimension. The general result is that adding dimensions helps, but less than we might have expected. The total information transmitted falls considerably short of a perfect addition of the information associated with the single dimensions. The largest number of dimensions was studied in an auditory experiment by Pollack and Ficks.[2] They used six dimensions and found 7.2 bits of information transmitted; this corresponds to the identification of about 150 different stimuli without error.

The limitations revealed by these studies seem quite severe. They clearly do not apply, as Miller reminds us, to those musically-sophisticated individuals with absolute pitch who can identify 60 frequencies. Miller also points out that the identifiability of faces, words, and objects in our ordinary behavior would suggest a far greater capacity than indicated by the studies of multidimensional judgments.

Can the limitations disclosed in these studies be overcome by means of a carefully contrived training procedure? In particular, will the procedures of automated instruction, or of teaching machines, enable us to exceed these limitations? The principal ingredients of automated instruction are (1) continual interrogation and overt response, (2) immediate feedback or knowledge of results, (3) learner-controlled pacing of the lesson, and (4) presentation of successive items conditional upon previous performance. These are the procedural principles to which almost all of the workers in the field of automated instruction would subscribe. A smaller, but substantial, number of them would support Skinner's addition of two more principles to this list: (5) gradual, "small-step" progression of the lesson, and (6) high probability of a correct response or positive reinforcement.[3] This paper reports an application of these procedures to the task of identifying multidimensional, nonverbal sounds.

Four experiments were conducted. In all cases the sounds were similar; they assumed two to five values along four or five dimensions, the dimensions being frequency, amplitude, interruption rate, duty cycle, and duration. The first two experiments examined the effects of the presence or absence of the lesson variables we have enumerated. On the basis of the results of these two experiments, and on the basis of our intuitions where gaps were present in the results, the third experiment was designed to maximize performance as measured by the information transmitted. In each of these three experiments, as in the experiment conducted by Pollack and Ficks,[2] the subjects identified the sound presented by listing the value assumed by each dimension, in short, by a four- or five-digit number. In the fourth experiment, certain of the procedures of automated instruction were re-examined under conditions in which the sounds were identified by arbitrary English words.

Apparatus and General Procedure

The stimuli were constructed and presented, and the lessons were monitored and analyzed, by a PDP-1B digital computer.[4] The subject sat at a typewriter connected to the computer, wearing earphones, in a remote room. Whenever he depressed the space bar, a trial began. Then, depending on the condition under study, a number of different things could happen. In some cases, the computer typed out the name of a sound and then presented that sound, after which the subject called for the next sound. In other cases, the computer presented a sound and then the subject attempted to type the correct name of the sound. After the subject's response, the computer might type OK or WRONG. After this feedback, the subject would call for the next sound, or, in some cases, if he were wrong, the computer might type the correct answer. In other instances the subject was allowed a second try, either immediately after the visual feedback, or after first hearing the correct sound again and also the sound corresponding to his response. In any case, on each trial the subject had only to depress those keys corresponding to the elements of his response; when appropriate the computer typed FADRL (for frequency, amplitude, duty cycle, repetition time, and length) on each trial, and it managed the necessary spacing and carriage returns for proper alignment of the subject's response.

Under some conditions, the sound on a given trial would be drawn at random, without replacement, from the entire set of sounds to be learned, and the whole set would be reinstated when it was exhausted. In another condition, the sounds withdrawn were replaced if not responded to correctly, so that each sound had to be correctly identified once before the entire set of sounds was reactivated. In some conditions, only a few of the sounds to be learned were in an active set, and these sounds were presented with the requirement that a certain level of mastery be reached before they were set aside in favor of another subset. At times these subsets were drawn at random; at other times they were selected to emphasize easy contrasts early in the lesson and more difficult contrasts as the lesson proceeded.

The computer constructed the sounds by setting up a sine table and reading the digitalized, time-quantized values (40 μsec step function) through a digital-to-analogue converter. A computer program was written to permit the experimenter to type in simply, at the beginning of a lesson, the sound parameters and the lesson characteristics that were desired. (Appendix B, p. 31, illustrates the process of setting up a lesson.) A time-sharing arrangement was used so that two subjects at their individual typewriters could be treated independently, within the general confines of a given lesson, at the same time.

Each experimental condition consisted of several one-hour training sessions. These sessions were conducted as nearly as possible on alternate days, with some exceptions made to accomodate the subjects' college class schedules. Four subjects served under each experimental condition. They were assigned to the various conditions of a given experiment on the basis of a simple pretest; the pretest required them to list the ways in which the members of several pairs of sounds differed from each other. Immediately after their training, all subjects were given an identification test under comparable conditions. For some subjects, this test was repeated after one month. The measures of performance secured from the training trials and the retention tests, though not all of them in all cases, were percent correct responses, latency of response, accuracy of confidence ratings, and number of bits of information transmitted.

Experiment I

The first experiment examined the effects of two types of response, of two types of feedback, of conditional and unconditional selection of successive items, and of three levels of reinforcement probability.

Specifically, a "standard" procedure (Condition 1) was compared with three "automated" procedures (Conditions 2, 3, and 4). Under the standard procedure, when the subject depressed the space bar the computer typed FADRL, effected a carriage return, identified a sound by typing the five-digit number that specified the value assumed by each of its dimensions, and then played the sound. The subject was passive

except for pressing the space bar to initiate the next presentation. Under Condition 2, when the subject pressed the space bar the computer typed FADRL, returned the carriage, and then played a sound; the subject typed the five-digit number he believed to identify the sound; the computer typed OK or WRONG and, if WRONG, typed the correct answer. Condition 3 was like Condition 2 except that on the trials on which he made an incorrect response on his first try, the subject listened to the sound again and then made a second try at identifying the sound correctly. Condition 4 was like Condition 3 except that when the first try was incorrect the subject heard the sound corresponding to his answer, as well as a repetition of the test sound, before making a second try.

Thus, under the standard procedure of Condition 1, the subject's response, if any, was covert, while under the automated procedures of Conditions 2, 3, and 4, the subject made an overt response. No feedback was given in Condition 1; visual feedback was given in Conditions 2 and 3; visual and auditory feedback were given in Condition 4. The selection of successive items in the first two conditions was independent of previous performance; in the last two conditions the order of presentation depended on performance to the extent that an item was re-presented immediately following an incorrect response. The varying structure of the different conditions led to three levels of reinforcement probability: this probability was zero under Condition 1, somewhat higher under Condition 2, and, because of the second try, still higher under Conditions 3 and 4. The expectation based on the principles of automated instruction enumerated above is that Condition 3 will be superior to Condition 2, which in turn will be superior to Condition 1. An expectation based on impressions obtained in some exploratory trials prior to the experiment is that Condition 4 will be superior to Condition 3.

On a given presentation each of the five dimensions of the sound could assume either of two values, so $2^5 = 32$ different sounds were presented. The two values were as different as they could be given the limitations of the apparatus, the range of sensitivity of the human ear, and time. The values of the temporal dimensions of the sounds were chosen insofar as possible to obtain an integral number of on-off cycles. The interaction between frequency and loudness was reduced by a correcting filter. No attempt was made to control the interaction between duty cycle and loudness. The sound parameters are listed in Appendix C, p. 39.

The sound on a given trial was drawn at random from the set of 32, and was not replaced. After each of the 32 sounds had been presented once (and some of them twice when a second try was called for), the entire set was reinstated. This process was repeated several times.

Four subjects served under each of the four conditions. The four groups of subjects were equated as nearly as possible on the basis of their pretest scores; each condition contained one subject from each of the quartiles of the pretest scores. After three one-hour training sessions, an identification test was administered. The identification test consisted of six presentations of each of the 32 sounds, a total of 192 presentations. The pacing of the test was controlled by the experimenter.

The first result to be considered is the probability of reinforcement that is obtained under each of the four conditions, since this variable serves as both a dependent and independent variable in the present experiment. As we have said, there were no overt responses, and hence there was a zero probability of reinforcement, in Condition 1. Defining a trial as including the second try, the average probability of reinforcement for the subjects in Condition 2 was approximately 0.70; the average reinforcement probability under both Conditions 3 and 4 was approximately 0.90.

The major results of this experiment, the percentages of correct responses on the identification test, are shown in Fig. 1. Consider first the means of the four conditions as denoted by $\bar{x}$'s in the figure. (Ignore for now the fifth condition shown there.) It can be seen that the standard procedure of Condition 1 -- with no overt response, no feedback, and zero probability of reinforcement -- led to the highest percentage of correct responses. Condition 4 -- with an overt response, an elaborate feedback, and a high probability of reinforcement -- led to the lowest percentage of correct responses. This difference is significant at the 0.05 level of confidence. Moreover, both Conditions 1 and 3 led to significantly better performance than Condition 2, at the 0.02 level of confidence. (Here, as elsewhere in this report, the significance tests discussed are t tests for the difference between correlated means, the correlation arising from the fact that the various groups were matched on the basis of the pretest.)

The analysis just described is based on 12 of the 16 subjects in the four conditions -- the subjects in the first three quartiles on the pretest. They are represented by filled circles in Fig. 1. The subjects who fell in the fourth quartile on the pretest are represented by open circles in the figure. It can be seen that in every case the fourth-quartile subject performed less well than the other subjects serving under the same experimental condition, and, in most cases, by a considerable amount. While recognizing the dangers of discarding subjects after the fact, we believe it advisable here to consider the alternative danger than an analysis based on all of the subjects may obscure some real differences among the four treatments. It may be the case that the pretest is a good indication of the extent to which the subject applied himself in the experiment. Given this possibility, we should probably

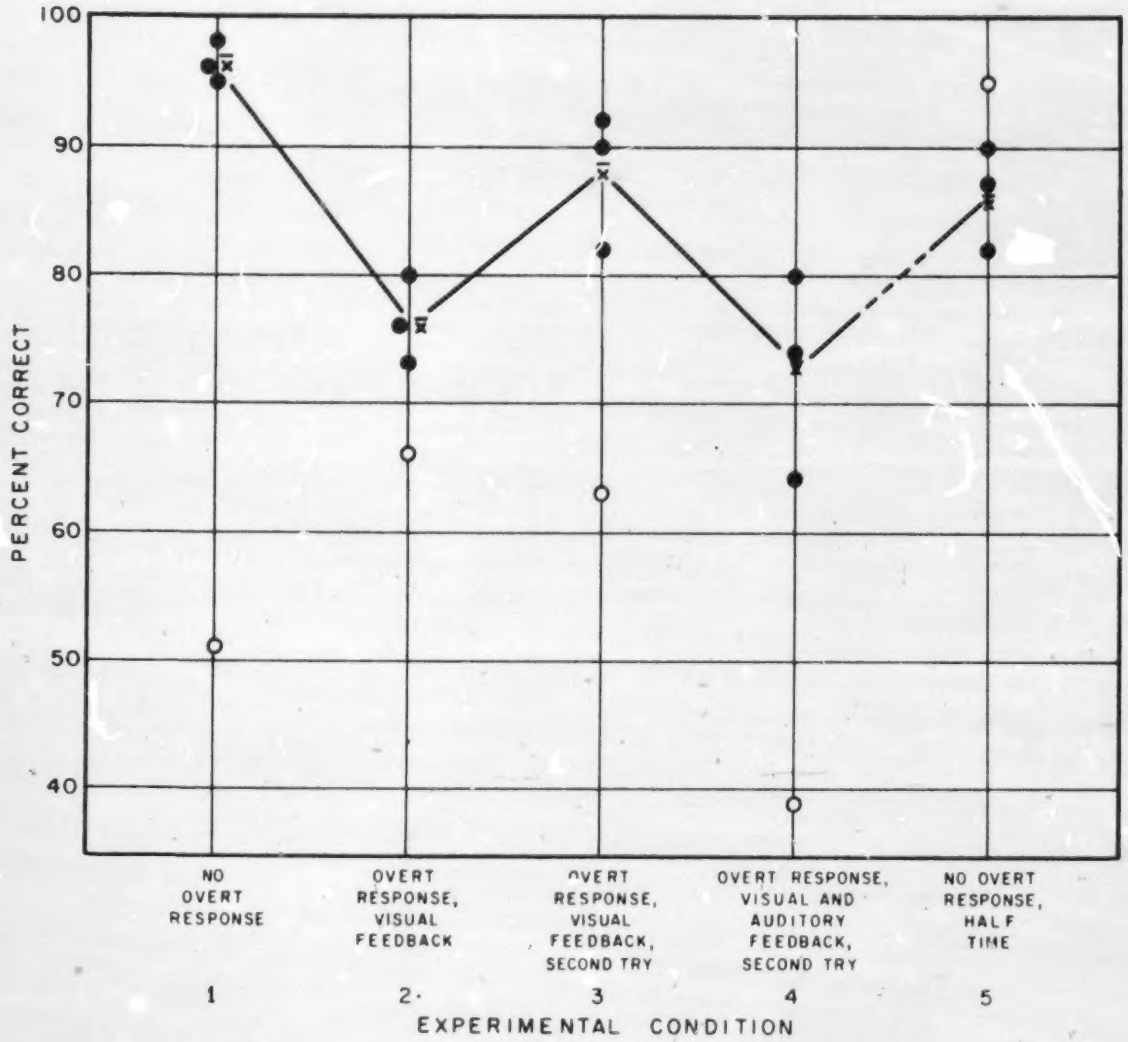


Figure 1. The Results of Experiment I

16

consider both analyses. An analysis based on all 16 subjects ranks the four treatments in the same way as the analysis based on 12; in this case the significant differences are those between Conditions 1 and 4 and between Conditions 3 and 4, at the 0.02 and 0.05 levels of confidence, respectively.

In any case, the results of this experiment were totally unexpected by, at least, the authors. They violate several of the major themes of automated instruction as well as the subjective impressions we gained while serving as subjects in some preliminary tests. It appears, however, that they may be rationalized to some extent.

In the first place, the finding that a covert-response technique is as effective or more effective than the overt-response techniques is in agreement, as it turns out, with several other studies. The field is young, so there are few published studies to cite, but a variety of personal communications indicates that this is coming to be a common finding. We can cite, by way of example, published abstracts of four meeting papers presented very recently.[5] Although we have said that there was no feedback or reinforcement in Condition 1, the subjects under that condition, if perverse, could have ignored the prompt until after the sound was presented. In that event, if we choose to define the term "response" to include covert activity, the subjects did receive feedback and, sometimes, reinforcement. Secondly, if we consider the procedure of Condition 1 as a kind of "prompting," as some investigators would, a superiority of Condition 1 over Condition 2 is in agreement with still other studies.[6] From this viewpoint, Condition 1 -- which we regarded as a "standard" procedure and, almost embarrassingly, as a straw man -- contains an important element of automated instruction. And, indeed, the procedure of Condition 2 is essentially the same as that used by Pollack and Ficks[2] before the influence

of automated instruction was felt. Finally, that Condition 4, in which auditory feedback was given, was relatively ineffective also seems reasonable when reconsidered. In this case the subject is provided an associated stimulus and response, the pair designated by his incorrect response, that intervenes between the elements of the stimulus-response pair that was supposed to be learned on that trial. This procedure thus produces a delay between the various stimuli and responses to be associated, and it may actively interfere with the association. Again, under this procedure, the subject may improve his initially incorrect response by concentrating on the difference between the two sounds presented, as originally intended by the experimenters. This process can, of course, deter the subject from building up a capability for absolute identification. This latter capability is the one desired, and also the one required by the identification test used as a measure of learning in this experiment.

Returning now to the presentation of experimental findings, the fact that an overt response was not required under Condition 1 led to the result that the subjects under this condition were exposed to more stimulus presentations than the subjects under the other conditions. The subjects under Conditions 2, 3, and 4 had an average of 250 trials while the subjects under Condition 1 received nearly twice that many, or 450, trials. In order to determine whether the sheer number of trials or the difference in procedure was critical in producing the results presented in Fig. 1, we examined a fifth condition which was identical to Condition 1 except that the time and the number of trials permitted were reduced by approximately one half. Before discussing the result let us note that, in one sense, we are asking a highly academic question. Practically, if one training procedure produces better results than another in the same amount of time, we would not discount the difference on the basis that the better procedure permitted more trials. Time, not number of trials,

is money. On the other hand, if it were established that the effectiveness of training is almost entirely dependent on the number of trials, this fact would constitute an important rule of thumb in the general design of training procedures.

The results of Condition 5 can be seen in Fig. 1. Again, four subjects are represented. They were selected on the basis of the pretest to obtain one subject from each of the quartiles as defined by the 16 subjects of the first four conditions. The mean shown, represented by $\bar{x}$, as in the case of the other means shown in the figure, does not include the subject from the fourth quartile. We see that, if the time is reduced by one half so that the number of trials is equated, the procedure not requiring an overt response is still relatively effective. Condition 5 is significantly better than Condition 2 at the 0.01 level of confidence, and not significantly different from the other conditions. We also see in the figure that the performance of the fourth-quartile subject in Condition 5 weakens somewhat the rationale for considering an analysis from which these subjects are excluded. However, as far as comparisons with Condition 5 are concerned, it makes little difference whether or not the fourth-quartile subject is included. If all subjects are included, Condition 5 is superior to Condition 2 at the 0.05 level of confidence and not significantly different from the other conditions. It seems clear, then, that the covert-response procedure is intrinsically more beneficial than a simple overt-response procedure. Perhaps the prompting is important, or perhaps having to make an overt response, in what is essentially a perceptual task, is disruptive.

To summarize the major results of this first experiment: (1) the data provide no support for the hypothesis that continual interrogation and overt response are beneficial to learning, (2) the data show that a fairly extensive feedback may be detrimental to learning, and (3) the data provide no support for the hypothesis that efficiency of learning varies directly with the probability of reinforcement.

Experiment II

The second experiment was concerned with the size of the step between successive presentations and with the dependence of the presentation on previous performance. Condition 1 can be termed a "whole method." The sounds were drawn at random without replacement from the entire set of sounds to be learned; when the set was exhausted the entire set was reinstated. The subject made an overt response, and received visual feedback. Condition 2 was a "whole method with dropouts." It was like Condition 1 except that the sounds responded to incorrectly were replaced, so each sound had to be correctly identified once before the entire set was reinstated. Condition 3 was a "part method." It was like the first two conditions in that the subject made an overt response and received visual feedback. It differed with respect to the selection of stimuli. A subset of two stimuli was chosen and the two were presented until both were correctly identified. Then another pair of stimuli was chosen and these two presented until both were correctly identified, and so on. The members of successive pairs were selected to form a gradual, "small-step," progression in terms of difficulty: the members of the first pair were as disparate as possible; the members of succeeding pairs were progressively more similar until in the last pairs the two stimuli differed by a single value on a single dimension.

Thus, under Condition 1, the successive presentations were independent of previous performance and it may be said that the step intervals were uncontrolled and, generally, large. Under Condtion 2 the presentation depended on previous performance, but again the step intervals were irregular in size and large. Under Condition 3 the presentation depended on previous performance and the step intervals were small. We would expect, on the basis of the principles of automated instruction that we have listed, that Condition 3 will be more effective than Condition 2, and that Condition 2 will be more effective than Condition 1.

The sounds in this experiment were composed of four dimensions; duty cycle was held constant, at 0.50%. Each dimension could assume three values, so $3^4 = 81$ different sounds were presented.

The three groups of four subjects were equated on the basis of pretest scores. Each of the 12 subjects chosen to participate in the experiment met a requirement that his pretest score be higher than the score that defined the boundary between the third and fourth quartiles in Experiment I. As in Experiment I, three one-hour training sessions preceded an identification test. The third session under Condition 3 differed from the first two sessions; the whole method was used in the third session to facilitate the transition to the identification test. The identification test contained 200 items.

The results of the experiment, in terms of the percentage of correct responses on the identification test, are shown in Fig. 2. The $\bar{x}$'s shown represent the mean result of each condition. It can be seen that the three procedures did not produce any significant differences. The trend is in the wrong direction. The data provide no support for the hypothesis that conditional presentations and small steps are conducive to learning.

Experiment III

Based on the results of the two experiments described above, on additional consideration of the principles of automated instruction, and on our intuitions, the training procedure of the third experiment was designed to produce maximal efficiency of learning. Compared with the earlier experiments, this one provided more extensive and varied training for fewer subjects on a larger ensemble of sounds.

Four subjects, three of them having served previously and relatively well in one of the first two experiments, were given seven hours of training on a set of 3125 sounds in which five dimensions could assume any of five values. In the first hour, only one of the five dimensions was varied in a given series of trials. The dimensions, then, were taken up singly; the order of presentation proceeded from the easiest of the five dimensions to the most difficult. Within the series of trials devoted to each dimension, the early trials presented the extreme values or the easier distinctions; later trials introduced gradually the intermediate values or the more difficult contrasts. The critical value presented on a given trial was typed out for

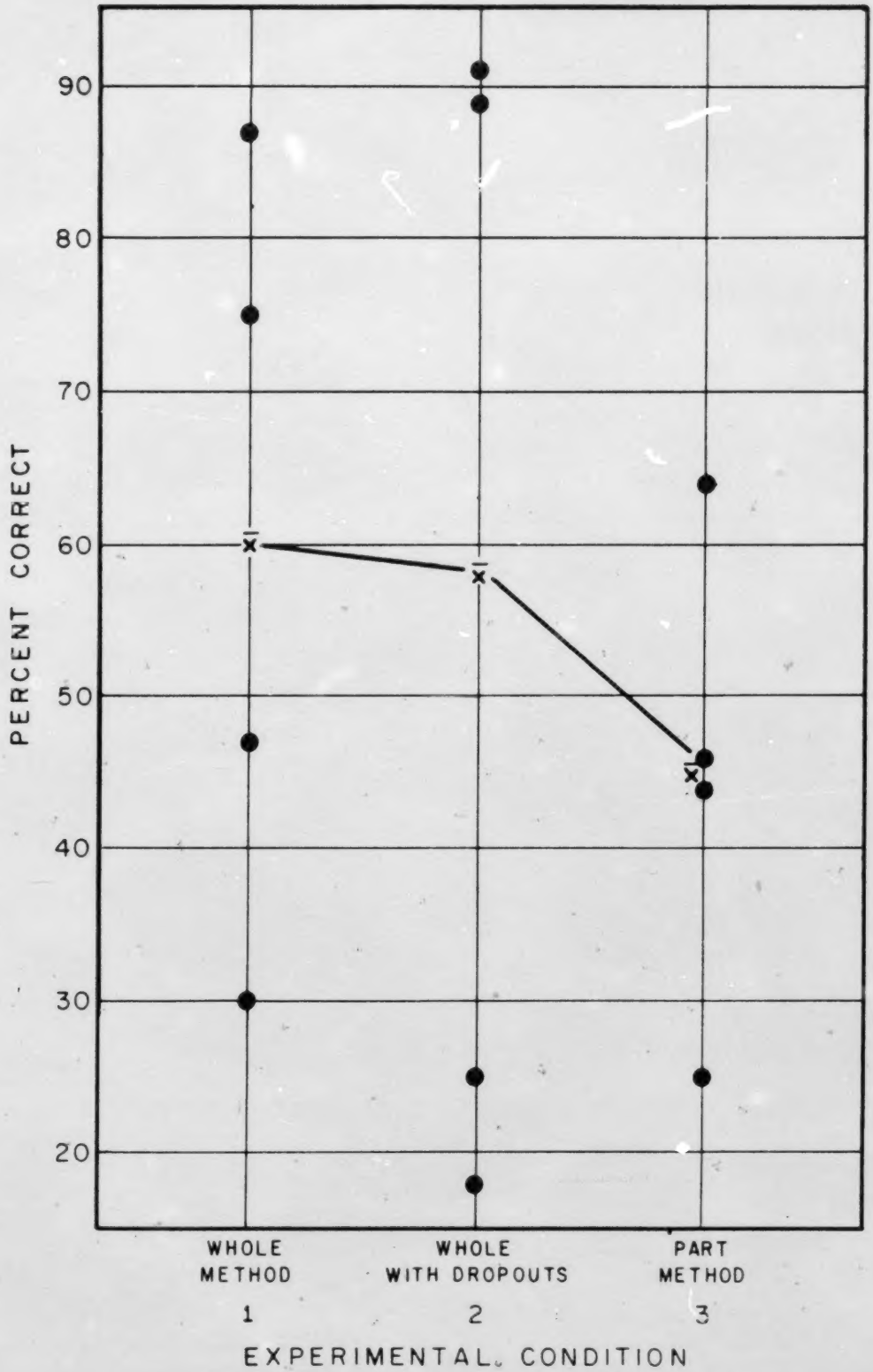


Figure 2. The Results of Experiment II

the subject before the sound was presented. The subject made an overt response with this prompt in view. In the second hour, the dimensions were again considered one at a time, but this time the values assumed by each dimension were presented in a random order, and no prompt was given. The subject made an overt response and received visual feedback. In the third, fourth, and fifth hours of practice, the sounds were drawn at random from the large set of 3125, and the subject responded to all five dimensions. The sounds were preceded by a prompt, but the prompt was removed from view before the response was to be made. In the sixth and seventh hours, random selections were again made from the large set; at this time the subject did not receive a prompt but did receive visual feedback.

The identification test in this experiment was paced by the subject. As a result, the number of items in the test varied slightly from subject to subject; the average number of items was 165. The essential results of the identification test are shown in Table 1. It can be seen that the average percentage of correct identifications is approximately 5%, and that the average number of bits transmitted was approximately 5.3.

The earlier results of Pollack and Ficks[2] provide a basis of comparison for these results. They found, it will be recalled, that an average of 7.2 bits of information were transmitted in their experiment. Although the figure of 5.3 bits just reported is lower, the adjustments necessitated by certain differences in the stimuli used in the two experiments lead to estimates of performance that are quite similar for the two experiments. In the study made by Pollack and Ficks, a sixth dimension was used (the direction of the sound) that was by far the easiest of the dimensions; moreover, the sounds in the earlier experiment varied in duration from 5 to 17 seconds, whereas in the present experiment they varied from 1 to 5 seconds. But the significant result, of course, is that our subjects did not perform substantially better than the subjects in the experiment conducted by Pollack and Ficks.

Table 1. The Results of Experiment III*

Subject	Percent Correct	Bits Transmitted
1	9	6.4
2	6	5.0
3	3	5.6
4	1	4.2

*With respect to the column headed "Percent Correct", a response was scored as correct only if the subject had listed the correct value for all five dimensions. The "Bits Transmitted", on the other hand, were calculated after the procedure of Pollack and Ficks[2]; the number of bits transmitted over the entire test was determined separately for each of the five dimensions, and these five numbers were summed to yield the figure shown for each subject.

Experiment IV

In the experiments described so far the subject was encouraged to develop an analytical attitude toward the sounds: he identified the sound by listing sequentially the value assumed by each dimension. Under this procedure the coding of stimuli and responses is very direct -- once having decided on the parameters of the sound presented, the subject has no doubts about the appropriate response. It is of interest to consider also a procedure in which the subject is encouraged, by the requirement of a unitary response, to register the sounds more immediately, and essentially as unitary events, as he does in the case of music or speech. This second procedure involves a less direct code -- the relationship of stimulus to response, as a matter of fact, may be completely arbitrary. It adds to the problem of identifying the sounds another distinct learning task, the task of learning "paired associates" as it is commonly called. In this case, the subject may be able to describe the sound in detail while being unable to produce its name. In terms of practical concerns, if our subjects can learn to respond synthetically to sounds of the type we have described above, we would expect the rate of information transmission (bits/sec) to be far greater than that attainable with the analytical response.

The sounds to be learned in the fourth experiment were 32 sounds drawn at random from the set of 625 sounds that correspond to five values on each of four dimensions. The sound parameters are those listed for Experiment III in Appendix C, p. 39, with the exception that the duration of the sound was held constant at 3 sec. Each of the 32 sounds was assigned an English word as its name. The names were words like "whale," "submarine," "helicopter," etc. -- objects in which our subjects might reasonably expect a sonar or aural radar operator to have some interest. A placard containing the 32 names was visible to the subject throughout the training sessions. The subject responded by typing only the first two letters of the name. He was not given any verbal information about the nature of the sounds; that is, he was not told that they were samples from a set of sounds with specifiable dimensions and dimensional values. As far as the subject knew, the sounds he heard represented the experimenter's conception of the way helicopters, etc., sounded to a sonar or radar operator.

Our principal concern in this experiment was to obtain an estimate of how well and how rapidly subjects could learn to identify sounds by their arbitrary names when trained under the procedures of automated instruction. We were also interested in a replication of some of the more striking comparisons of the previous experiments, partly because the surprise value of those comparisons suggested a replication, and partly to determine whether or not the earlier results would also be obtained with the different learning task of the present experiment. This experiment, therefore, compares a procedure under which prompts are given and no overt response is made with a procedure in which an overt response is made and visual feedback is given, the presentation of stimuli in both instances being that of the whole method. These procedures correspond essentially to Conditions 1 and 2 of Experiment I and Fig. 1. A third procedure was like the second except that the part method was used; whenever each of a subset of four items was correctly identified, another subset was presented. Thus, the second and third procedures of this experiment reproduce the essentials of Condition 1 and 3 of Experiment II and Fig. 2.

Four subjects were assigned to each of the three conditions. The performance of the subjects on the pretest met the criterion of adequacy as defined in Experiment I. They were given four hours of training. The identification test consisted of approximately 200 items.

The results are shown in Fig. 3. It can be seen, first, that the most effective procedure led to approximately 35% correct identifications. Although no comparisons are available, this result does not seem to provide a great deal of support for the use of automated procedures in auditory learning. We could hardly term it a "breakthrough."

It can also be seen in the figure that the prompt-plus-covert response procedure yielded a higher percentage of correct identifications than the overt response-plus-feedback procedure -- as was the case in Experiment I. And the whole method yielded a higher percent correct than the part method -- as was the case in Experiment II. The differences between Conditions 1 and 2 and between Conditions 1 and 3 are significant at the 0.02 and 0.05 levels of confidence, respectively.[7]

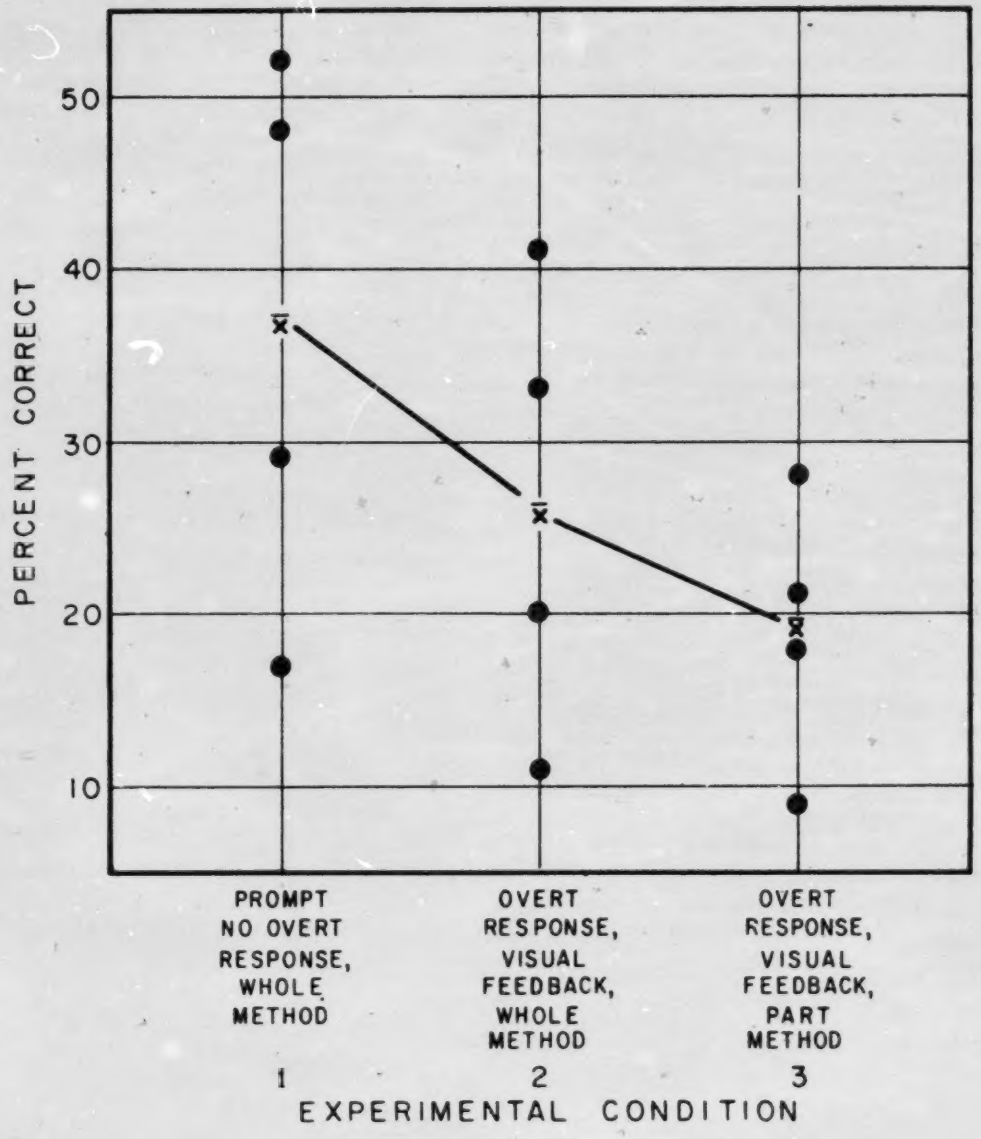


Figure 3. The Results of Experiment III

Secondary Measures of Performance

The latency of the response was measured in each of the experiments, and confidence ratings were obtained in the fourth experiment. Our purpose was to examine the possibility that these two measures could be used to advantage in determining the order of presentation of stimuli, or, in other words, in making subsequent presentations appropriately conditional upon previous performance. The major results of the experiments conducted to date have not encouraged us to analyze these measures in any detail. We can say, however, that the distributions of these measures were different for correct and incorrect responses. In Experiment IV, for example, the difference between the mean latencies associated with correct and incorrect responses was significant at the 0.01 level of confidence, and the difference between the mean confidence ratings was significant at the 0.001 level of confidence.

The measures of retention taken at the end of one month deserve more discussion. The subjects in Experiment I were given an identification test after one month identical to that given immediately after the training. Almost no loss was observed. For Condition 1, the average identification score was 97% of the score secured immediately after training; for Condition 2, 101%; for Condition 3, 99%; and for Condition 4, 83%. It is probably significant that the subjects under Condition 4, who performed least well in the original sessions, suffered the most from the lapse of time.

The one-month retention test administered to the subjects of Experiment IV, who had learned to identify the sounds by English words, employed the "recognition-memory" procedure. Under this procedure the stimuli to which the subject had been exposed previously are mixed with a set of new stimuli, and the subject is required to state his confidence that the stimulus on a given presentation was a member of the original set. The measure of sensitivity d', frequently used in psychophysics because it is independent of the subject's response bias, can be computed from the resulting data and interpreted as a measure of retention. It will be recognized that this procedure is a very sensitive one: it taps residual memory inadequate to yield correct absolute identifications.

We shall not describe the analysis procedure in any detail, since it has been presented previously.[8] The major result is that the subjects of Experiment IV retained very little of what they had learned. The average value of $\underline{d}'$ was 0.30; this figure corresponds to 55% correct responses in a yes-no test in which the chance probability is 0.50. A test of absolute identification given at the same time yielded a consistent result: the average percent correct responses dropped, from approximately 25% at the time of learning, to approximately 10% one month later.

Given the results of Experiments I and IV and the retention scores of the subjects in these experiments, it seems clear that a set of sounds of a given size is much easier learned and very much better retained when they are analyzed and identified by a direct numerical code that describes their properties than when they are taken as units and identified by some arbitrary label.

Conclusion and Discussion

The overwhelming result of this research is that the techniques of automated instruction, in our hands, have failed to produce high rates of learning. How should we qualify the conclusion that the automated training procedures are little, if any, more effective than conventional procedures?

We would certainly restrict this conclusion to perceptual learning, and perhaps to auditory learning. The negative conclusion might not apply to tasks presented to another sensory modality, and there is no basis for including in its scope tasks having a greater cognitive content.

Moreover, there are two factors that might conceivably mitigate the conclusion as it applies to auditory learning. One is the problem of the motivation of experimental subjects. The second is the problem of translating procedural ideas into actual experimental procedures. Both of these issues, of course, are general to a wide class of experiments.

It may be the case, as many suspect, that the procedures of automated instruction, to the extent that they are effective, do not make learning more effective for highly-motivated students, but instead affect learning by eliciting sustained performance from students with little independent motivation for the particular task posed. It may also be the case that our subjects approached the task we posed for them with a high level of motivation. It can be argued that, in general, experimental subjects respond with a highly-motivated performance simply because they are experimental subjects, because someone is paying attention to them -- the "Hawthorne effect" of social psychology. If motivation was thus assured for the subjects in our experiments on grounds essentially independent of the task we gave them, we would expect differences among the various automated procedures, and differences between automated procedures as a class and the standard procedures, to be minimized.

Again, it may be the case that our results do not bear upon basic principles of learning, or auditory learning, but reflect instead inequities created when the ideas about procedure were physically realized. We have not been able, of course, to convey to the reader in words a real sense of the ease or difficulty of the mode of response, the felicity of the different kinds of feedback, or the propriety of the various timings of events. These matters might well be critical. If the response, or feedback, or timing in our experiments were awkward, the automated procedures would suffer in comparison. It might be, too, that however well designed these features are, the mere need for frequent motor responses and the relentless supply of feedback information will be actually disruptive for the highly-motivated person. It is true that the most effective of the procedures we considered, the standard procedure, was also the most streamlined procedure.

In our opinion, these two general problems -- of subject motivation and experimental realization -- do not apply with any special strength to the experiments that we have presented; they do not seem to us to require specific qualification of the results or conclusions of these experiments. With respect to motivation, it may be noted that the skills we offered to teach our subjects were not of any value to them, and that we were clearly unable to influence their lives in any way beyond the four or five hours they

spent in the laboratory; thus, two usual incentives for motivated performance were lacking in these experiments. It is conceivable that a response more convenient than typing might be found, for example calling out the response, or pushing buttons arrayed in a dimension-by-value matrix, but we doubt that such modifications would have a substantial effect. None of the various automated procedures appeared to us to be particularly inconvenient, and the subjects did not voice complaints on this score. We take the position, in short, that the major result of this investigation is likely to be valid for auditory identification. At the very least, it must be granted that the principles of automated instruction are not ready for cavalier application in psychoacoustics.

REFERENCES

1. G. A. Miller, Psychol. Rev. 63, 81-97 (1956).

2. I. Pollack and L. Ficks, J. Acoust. Soc. Am. 26, 155-158 (1954).

3. B. F. Skinner, Science 128, 969-977 (1958).

4. Digital Equipment Corporation, Maynard, Massachusetts. (See Appendix A, p. 26, for a description of the computer.)

5. See American Psychologist 16 (1961): F. F. Kopstein p. 466; H. H. Shettel and R. H. Lindley, p. 369; A. Roe, p. 369; E. R. Keisler, p. 369.

6. J. O. Cook and T. S. Kendler, In G. Finch and F. Cameron (eds.) Symposium on Air Force human engineering, personnel, and training research, Nat. Acad. Sci.-Nat. Res. Coun. Public. 455, 90-98 (1956); J. O. Cook, J. exp. Psychol. 56, 455 (1958); J. O. Cook and M. E. Spitzer, J. exp. Psychol. 59, 275-276 (1960); F. F. Kopstein and S. M. Roshal, Amer. Psychologist 10, 354 (A) (1955). These studies were pointed out to us by F. F. Kopstein.

7. Observation of Fig. 3 may lead the reader to wonder if the differences reported here as statistically significant really are. It may also appear strange that the difference between Conditions 1 and 2 is reported to reach a greater level of confidence than the difference between Conditions 1 and 3, in view of the fact that the difference between 1 and 3 is greater than the difference between 1 and 2, and the variance of 3 is less than the variance of 2. It was mentioned earlier that, because these experiments employed matched groups, the formula for the difference between correlated means was used in the computation of t. The use of matched groups ordinarily reduces the sampling error, and, in these experiments, did so appreciably. The extent of the reduction, of course, depends on the extent to which the variable used in matching is correlated with the variable under study. The correlation was higher for Conditions 1 and 2 than for Condition 3.

8. J. P. Egan, "Recognition Memory and the Operating Characteristic," Tech. Note AFCRC-TN-58-54, Hearing and Communication Laboratory, Indiana University (1958). I. Pollack, J. Acoust. Soc. Am. 31, 1126-1128 (1959).

APPENDIX A

The PDP-1B Computer

The Programmed Data Processor (PDP-1B) is a high spee solid state, digital computer designed to operate with several types of input-output devices, with no internal machine changes. It is a single address, single instruction, stored program computer. Five-megacycle circuits, a magnetic core memory, and fully parallel processing make possible a computation rate of 100,000 additions per second.

It has a random-access, magnetic-core memory consisting of 4096 18-bit words. This memory has a five-microsecond cycle time.

Of particular interest is the ease with which new and unusual external equipment can be added. For example, the computer is capable of handling more than 100 typewriters simultaneously -- each typewriter operating independently.

Input-Output Equipment

Typewriters

The two Soroban typewriters installed on PDP-1B are used as both input and output devices. They are, essentially, 44 key IBM electric typewriters modified for use with a computer.

Keyboard

A small keyboard to facilitate the entry of numerical data is installed on the computer.

Visual Display

A most versatile output unit is the 19-inch cathode-ray tube display. This is essentially a very-high-speed x-y point plotter. It plots points at a rate of 20,000 per second. By proper programming it is possible to display letters, sentences, graphs, charts, pictures, etc. These displays may be viewed as they are produced, or recorded on film for later reference.

Light Pen

The light pen is an input device that is used in connection with the visual display. It may be used to point out to the computer an item in the display that is of interest to the operator, or it may be used to allow the input of freehand drawings -- of graphs, forms, block printed letters, etc.

Paper Tape Readers

The photo-electric paper tape reader is used as the major input for programs and stored data. It is capable of reading any standard-width paper tape at a rate of 400 characters per second.

Paper Tape Punches

There are two paper tape punches mounted on the computer. These punches are capable of punching any standard-width paper tape at a rate of 60 characters per second. One punch and one typewriter may be interconnected so that tapes may be prepared without involving the computer.

Digital-to-Analog Converters

There are 3 digital-to-analog converters on the computer. These converters may be used to change numerical data in the computer into a voltage function of time. This facility is useful in providing a simple way of generating audio signals and in controlling electronic apparatus.

Analog-to-Digital Converter

An Epsco Transicon analog-to-digital converter converts voltages into numbers. It is able to commutate several signal sources. A typical use is to convert electroencephalographic signals into numbers and thereby to allow analysis of these data by the computer. In this connection, a program has been written for PDP-1B to simulate the ARC (Average Response Computer), and the program has been used to analyze the EEG response evoked by an auditory click.

Clock

To permit the computer to measure time intervals accurately, there is a "clock" that can be read by the computer. The clock also allows the computer to determine the time of day.

Utility Programs

Programs currently available in our computer facility are:

Typewriter Control Program

This program, now in use, allows a computer user to examine the contents of those memory locations, to search for particular contents, etc. Its main feature is that it allows remotely located users to control and debug programs.

FRAP

This is a symbolic assembly program. It allows the user to write programs with variable-length symbols rather than binary or octal numbers.

DECAL

This is a more complete and elaborate assembly-compiler program. It allows the definition and use of algebraic expressions, and, in general, it provides a means of defining a language oriented about a particular problem. It also accepts the statement of a procedure to solve a problem in this newly defined language. It then converts the problem-oriented-language statement into machine code in a form ready to run in the computer. This program is similar to, but more general than, the IBM FORTRAN System.

Floating-Point Subroutines

A complete package of floating-point subroutines for numerical computation. These include all of the standard trigonometric functions, logarithm, exponential, square root, number conversion, etc.

Statistical Programs

These are programs for the solution of basic statistical problems. They include mean and standard deviation, correlation, histogram generation, etc.

Display Routines

These routines display graphs of various types, e.g. continuous functions, point graphs, bar graphs, histograms, X-Y-Z plots (such as amplitude-frequency-time plots with scope brightness representing the Z variables). The routines also display coordinate systems, e.g., log-log, semi-log. There are facilities for recording scope plots on film as a permanent record.

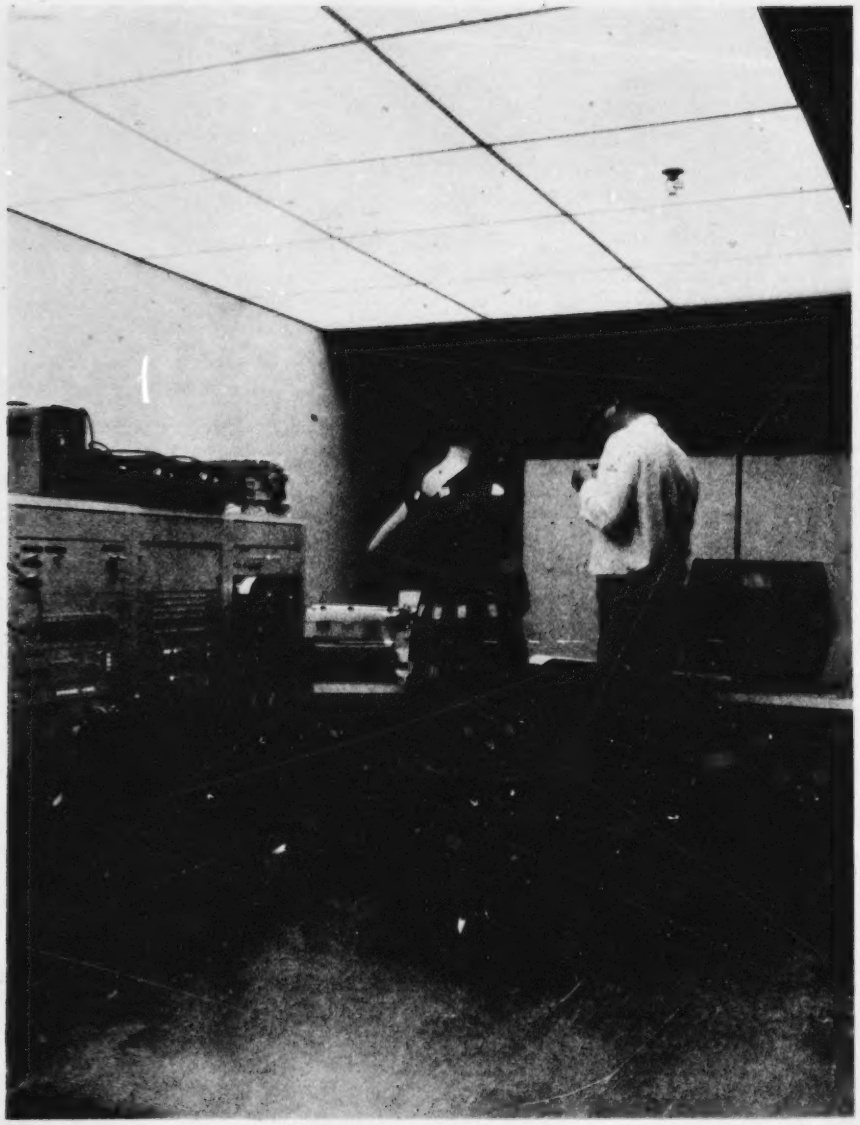


Figure A-1. Photograph of the PDP-1B Computer

APPENDIX B

Programming and Lesson Procedures

In the following we describe the use of the computer program that has been written to make a teaching machine of our computer. These pages illustrate the process of setting up an experimental training session on the computer, the course of a simple lesson, and the error-checking features of the program.

During an experimental session, or lesson, a set of sounds is formed and played to a subject. The contents of this set of sounds and the conditions of its presentation are controlled by the experimenter prior to the lesson. The experimenter is aided by the computer in determining the sound parameters and the values of the other lesson variables for a particular lesson; the computer asks him in a series of questions to specify these. The questions asked by the computer are shown in Fig. B-1.

When the program is loaded into the computer it stores a table of sound parameters. (In particular, the program specifies five alternative values for each of the five dimensions of the sound: frequency, amplitude, repetition time, duty cycle, and length.) Depending upon the circumstances, the experimenter may wish to use these parameters or insert new ones. Several sound-parameter tapes are available for different purposes. The experimenter may read in one of these, or he may insert new parameters directly from the typewriter.

When the lesson begins, the entire set of sounds, as determined by the parameters, is considered "active" and token sounds are picked from it at random. The experimenter designates a value "K" which states how many times a sound must be correctly responded to before it is dropped from the active set. When the subject reaches the criterion "K" for a particular sound, it becomes inactive, i.e., it is not presented again, until every sound in the original set reaches the criterion. When all the sounds reach the criterion, the entire set may be reactivated so that each sound must again reach the criterion. The number of times the procedure is performed is controlled by the experimenter by designating a value "R".

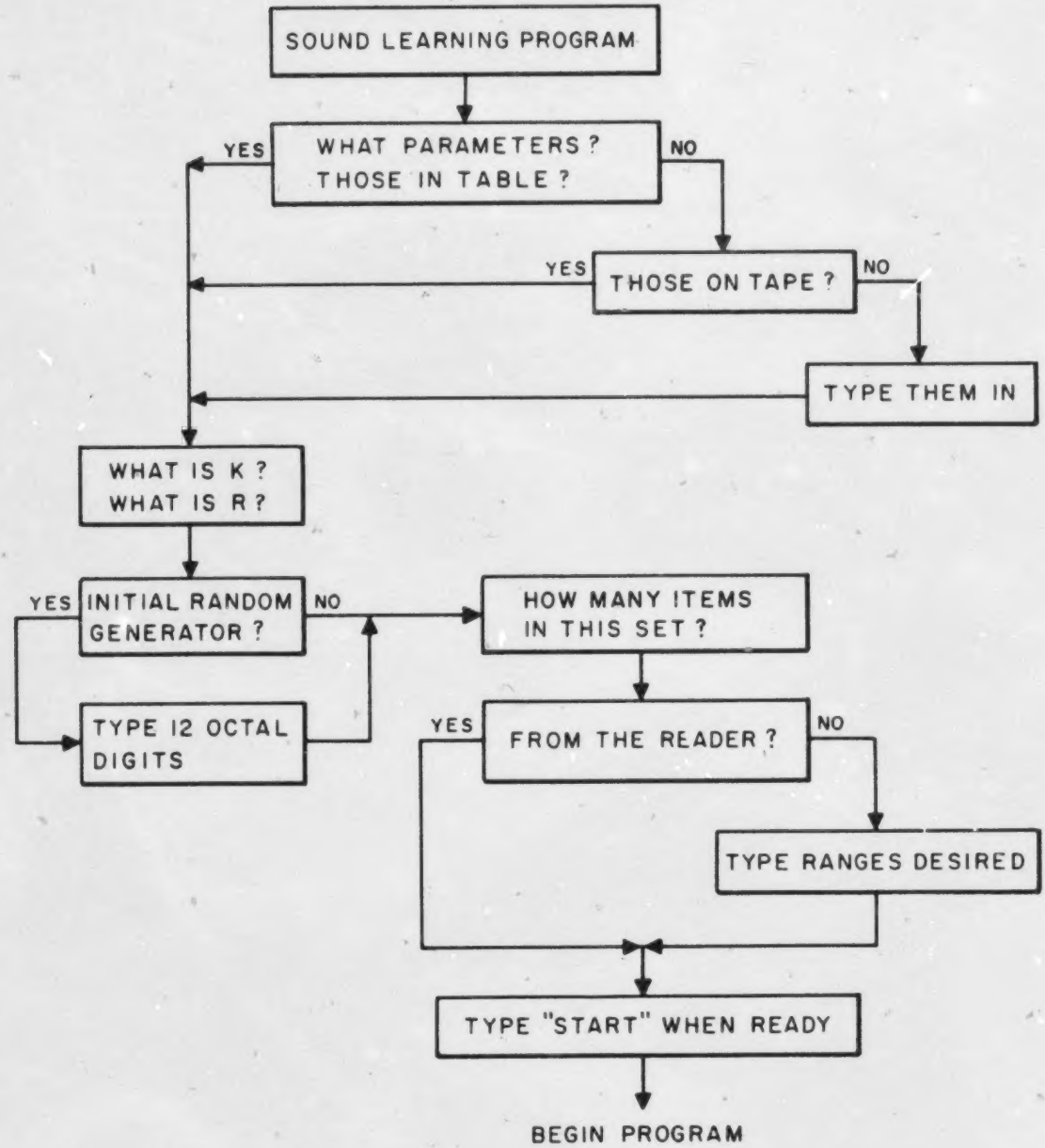


Figure B-1. Flow Chart of an Illustrative Computer Program

The experimenter must indicate how many sounds the present lesson will contain, that is, it may contain fewer sounds than provided for by all the possible combinations of the sound parameters. He may read in on tape the specific sounds he desires to use, or he may specify the ranges of the parameters from which the computer is to choose (at random) the indicated number of sounds. If the experimenter wishes to present the same "random" selection of sounds that was used in some previous lesson he may respond to an earlier question in such a way as to achieve this: he "initializes the random generator" by typing a particular set of 12 octal digits.

At this point the lesson is completely spelled out, and the experimenter calls in the subject. The computer has printed out for the subject the instruction to "TYPE 'START' WHEN READY TO BEGIN." When the subject types "START," the computer types out some additional instructions specific to the lesson at hand before beginning to present the sounds to the subject. The subject will have had previously some general instructions and some practice sessions.

The following examples (1-6) were produced by the computer program. They show the questions asked by the computer and the response by the experimenter or subject. Examples 1, 2, and 3 show the various ways of setting up a lesson. They illustrate the ease with which the sound parameters and the other lesson variables may be changed, thereby giving the experimenter a large amount of flexibility.

In choosing sound parameters to present to the subject, the experimenter can choose from the following ranges: frequency 100-8333 cps, amplitude 0-40 db (above an adjustable reference SPL), duty cycle 0-100 per cent, repetition time 0-60,000 milliseconds, and length 0-60,000 milliseconds.

Example 1 shows new sound parameters being entered into the computer by means of the typewriter. In this example each stimulus must be responded to correctly twice (K=2) before it is removed from the active set. The entire set is presented to the student only once (R=1). In the example shown, the sound presented to the subject may take on any of the five designated frequencies, any of the five amplitudes, a duty cycle of 55 per cent, a repetition time of 300 milliseconds, and a length of 2000 milliseconds. These values are designated

EXAMPLE 1

```
              SOUND LEARNING PROGRAM
WHAT SOUND PARAMETERS?
              THOSE IN THE TABLE? NO
              THOSE ON THE TAPE? NO
THEN TYPE THEM IN.

              FREQUENCY AMPLITUDE DUTY-CYCLE REP-TIME LENGTH
              200       10        20         100      500
              300       20        40         200      1000
              400       25        55         300      1500
              800       32        75         400      2000
              1600      40        90         500      2500

WHAT IS "K"? 2
WHAT IS "R"? 1
INITIALIZE RANDOM GENERATOR?NO
HOW MANY ITEMS IN THIS SET?5
FROM THE READER?NO
I WILL GENERATE THEM.  TYPE RANGES DESIRED.
              FREQUENCY AMPLITUDE DUTY-CYCLE REP-TIME LENGTH
              1-5       1-5       3-3        3-3      4-4
TYPE "START" WHEN READY TO BEGIN.
```

EXAMPLE 2

```
              SOUND LEARNING PROGRAM
WHAT SOUND PARAMETERS?
              THOSE IN THE TABLE? NO
              THOSE ON THE TAPE? YES

WHAT IS "K"? 2
WHAT IS "R"? 1
INITIALIZE RANDOM GENERATOR?NOHOW MANY ITEMS IN THIS SET?5
FROM THE READER?YES
TYPE "START" WHEN READY TO BEGIN.
```

EXAMPLE 3

```
              SOUND LEARNING PROGRAM
WHAT SOUND PARAMETERS?
              THOSE IN THE TABLE? YES
WHAT IS "K"? 2
WHAT IS "R"? 1
INITIALIZE RANDOM GENERATOR?NO
HOW MANY ITEMS IN THIS SET?5
FROM THE READER?YES
TYPE "START" WHEN READY TO BEGIN.
```

by typing the ranges desired from the parameter table. In Example 2 the experimenter read into the program a tape containing the desired parameters, and also a previously prepared tape containing the desired ranges of these parameters. Example 3 shows the case where the parameters stored on the original program tape are to be used. As before, a tape is read in designating the ranges desired of these parameters.

Example 4 is a simple lesson in which the student is to identify a sound which may vary only in frequency. When the student types "START," instructions are typed out to him by the computer. These instructions may be changed from student to student or from time to time depending upon particular requirements. The lesson then continues at a pace set by the student until the lesson is completed. In the sample shown the lesson proceeds as follows: when the student is ready he types "START." The computer then types instructions to the student. Then it types the letters FADRL. These letters stand for frequency, amplitude, duty-cycle, repetition-time, and length. When the student feels he understands the instructions and is ready to begin the lesson he depresses the space bar. This action produces in his headset the sound he is to identify. He then types his answer in the following way: Assume the frequencies that could be presented are those shown in Example 1. If he believes the sound presented was a 200 cps sound he types a "1"; if it was a 300 cps tone he types a "2"; etc. The same would be true for the other dimensions. In this example, however, these other dimensions are fixed at 3434 for A, D, R and L respectively so the subject types his choice of frequency and then 3434. If his response is correct the computer types "OK," effects a carriage return, and awaits the call (a depression of the space bar) for the next sound. If however a wrong answer is given, the computer types "WRONG," reproduces the sound corresponding to the subject's response, and then the original sound. Then it asks the subject if he wants to try again. If he types "YES," the computer repeats the above. If he types "NO," the computer types the correct answer, repeats the sound and asks the subject to type the correct answer before proceeding to the next sound. This entire process is continued to the completion of the lesson.

EXAMPLE 4

```
START
THIS TYPES INSTRUCTIONS TO THE SUBJECT.
              FADRL
              13434 OK
              FADRL
              33434 WRONG  TRY AGAIN?YES
              FADRL
              43434 OK
              FADRL
              53344 WRONG  TRY AGAIN?NO
CORRECT ANSWER IS....53434
              FADRL
              53434 OK
              FADRL
              23434 OK
              FADRL
              23434 WRONG  TRY AGAIN?YES
              FADRL
              33434 OK
              FADRL
              13434 OK
              FADRL
              23434 OK
              FADRL
              33434 WRONG  TRY AGAIN?NO
CORRECT ANSWER IS....53434
              FADRL
              53434 OK
              FADRL
              33434 OK
              FADRL
              43434 OK
LESSON COMPLETED
```

NOTE: Since this example was prepared, the computer program was modified so that the subject will have to respond only to those dimensions that are being varied in his lesson. In the example shown this means he would only be asked to type a number for the frequency. We may also note that the amount and kind of feedback information is variable, i.e., in some experiments the computer will not reproduce the sound corresponding to the subject's response, etc. Again, the subject may not be given the option of whether to try again after a wrong response. Another feature incorporated since this example was prepared is the use of red as well as black type; in general, the material typed by the computer will be one color and that typed by the subject another.

Examples 5 and 6 show the error-checking features of the program. In Example 5, line 4, the computer asked "THOSE ON TAPE?" The experimenter typed "YES" and then attempted to read in a wrong tape. The computer rejected this, typed "WRONG TAPE" and asked the question again. The experimenter then typed the letter "U." This was an inadmissible answer so the computer rejected it, instructed the experimenter to "TYPE 'YES or NO'," and asked the question again. The same thing occurred in Example 6, line 2, where the subject typed an inadmissible answer. In line 8, the computer asked "WHAT IS R?" The experimenter responded by typing an alphabetic character. In this case the computer wanted a numeric character so it rejected the response, typed "ERROR" and waited for a new response. This process continued until a numeric answer was given. This occurred also in Example 5, line 17, and Example 6, line 4 and line 8. In Example 5, line 10, the computer asked for 12 octal digits. "Eight" is not octal so it was rejected. This continued until 12 octal digits, not all of which were zero, were typed in. As is seen, the computer will not advance until the right kind of information has been given it. If the wrong kind of information is presented, the computer will reject it, instruct the user about the error, and then ask for the information again.

EXAMPLE 5

```
                SOUND LEARNING PROGRAM
WHAT SOUND PARAMETERS?
                THOSE IN THE TABLE?  NO
                THOSE ON THE TAPE? YES  WRONG TAPE
                THOSE ON THE TAPE? U TYPE "YES"  OR "NO".
                THOSE ON THE TAPE? YES
WHAT IS "K"? 2
WHAT IS "R" AERROR 1
INITIALIZE RANDOM GENERATOR? YES
TYPE TWELVE OCTAL DIGITS. 8
TYPE TWELVE OCTAL DIGITS. 000000000000NOT ALL ZEROES.
TYPE TWELVE OCTAL DIGITS. 123456712345
HOW MANY ITEMS IN THIS SET? 5
FROM THE READER? NO
I WILL GENERATE THEM.  TYPE RANGES DESIRED.
                 FREQUENCY AMPLITUDE DUTY-CYCLE REP-TIME LENGTH
                 1-BERROR
                 1-5       3-3       4-4         3-3     4-4
TYPE "START" WHEN READY TO BEGIN.
STRERROR
```

EXAMPLE 6

```
                FADRL
                55555 WRONG  TRY AGAIN?YER TYPE  "YES" OR "NO".YES
                FADRL
                RERROR
                34343 WRONG  TRY AGAIN?NO
CORRECT ANSWER IS....53434
                FADRL
                RERROR
                53434 OK
                FADRL
                13434 OK
                FADRL
```

APPENDIX C

Sound Parameters

The values given are (F) cps, (A) db re 0.0002 d/cm^2, (D) % on time, (R) on-off cycle in millisec, (L) millisec.

Experiment I

	F	A	D	R	L
1.	694	60	10	500	2000
2.	6250	95	90	250	4000

Experiment II

	F	A	D	R	L
1.	694	60	-	1000	2000
2.	1785	75	50	500	3000
3.	6250	95	-	250	4000

Experiment III

	F	A	D	R	L
1.	510	55	10	1000	1000
2.	1250	65	30	800	2000
3.	2500	75	50	500	3000
4.	4166	85	70	400	4000
5.	6250	95	90	200	5000

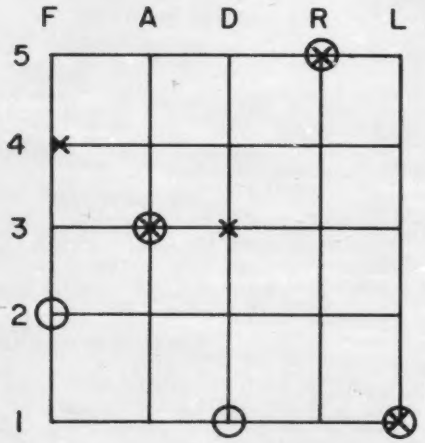
Experiment IV

The sound parameters were identical to those of Experiment III except that L was held constant at 3000 millisec.

APPENDIX D

Recommendations for Further Research

1. As mentioned in the BRIEF OF THE STUDY, it is conceivable that a less cumbersome mode of response, and a more natural form of feedback, would make the automated procedures more effective. Whether or not our predictions are optimistic on this score, the possibility should be investigated. One way to achieve what might be a more convenient stimulus-response coding would be to use a scope and light pen as the computer's in-out device. A grid can be projected on the screen with the intersections representing the several values of the several dimensions, thus:



(where F = frequency

A = amplitude

D = duty cycle

R = repetition time, and

L = length).

The subject would respond by pointing to a series of intersections with the light pen. The pointing would leave symbols on the screen representing the subject's identification of the sound (see x's in the figure); this answer is then compared with the correct answer (feedback) projected on the screen after his response (see circles in figure). The coding under this arrangement is very direct. A high pitch now coincides not only with a high number, but is also high in the spatial representation. The learner could see at a glance, in the example given, without translation through a numerical code, that he had identified A, R, and L correctly, and had underestimated F and D.

2. Learning situations should be constructed in which the subject has greater freedom in approaching his task. Our subjects followed rigid methods of study that we dictated. The computer as a teaching device is probably best exploited if use is made of its flexibility. The subject might be given the ability to produce any sound he chooses at any time, that is, he might be given the ability to interrogate the device. Then, if he believes a certain dimension is giving him special difficulty, he can explore that dimension in any fashion he chooses for as long as he chooses. In short, he can study or practice in his preferred pattern. He would also be allowed, under this scheme, to attempt a test at his discretion. He could, further, be allowed to determine the aspect of the subject matter to be emphasized in a test, and the nature of the test, for example, whether it will include prompts, or feedback, or provision for second tries when incorrect.

3. The possibility should be examined that successive items can be selected to be conditional in a more subtle way on previous performance. In the experiments reported here, only the frequency of correct responses on the particular sounds determined the next item. The selection might be made to depend on the combined basis of the latency of the responses, the confidence ratings attached to the responses, and the magnitude of the errors. Good indices of the subject's state of knowledge might be used to advantage.

4. Although it will be difficult to accomplish, an attempt should be made to relate the effectiveness of different procedures to individual differences among students. This study would involve an examination of the size of the step interval between successive items, and of the kind of response and feedback. In general, an examination of lesson formats with various degrees of streamlining should be studied in connection with individual characteristics of the learners.